UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/658,312	09/09/2003	I.V. Ramakrishnan	8009-23	2158

22150   7590   11/15/2006

F. CHAU & ASSOCIATES, LLC
130 WOODBURY ROAD
WOODBURY, NY 11797

EXAMINER
MYINT, DENNIS Y

ART UNIT	PAPER NUMBER
2162	

DATE MAILED: 11/15/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

	Application No.	Applicant(s)
Office Action Summary	10/658,312	RAMAKRISHNAN ET AL.
	Examiner	Art Unit	
	Dennis Myint	2162	•

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on <u>17 October 2006</u>.

2a)☐ This action is **FINAL**.   2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) <u>1-16 and 19-22</u> is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) <u>1-16 and 19-22</u> is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on <u>09 September 2003</u> is/are: a)☒ accepted or b)☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a)☐ All   b)☐ Some * c)☐ None of:

      1.☐ Certified copies of the priority documents have been received.

      2.☐ Certified copies of the priority documents have been received in Application No. _____.

      3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1)☐ Notice of References Cited (PTO-892)

2)☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3)☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
    Paper No(s)/Mail Date _____.

4)☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date _____.

5)☐ Notice of Informal Patent Application (PTO-152)

6)☐ Other: _____.

DETAILED ACTION

1.   This communication is responsive to Applicant's Amendment, filed on 17 October

2006.

Response to Arguments

2.   Applicant's arguments filed on 17 October 2006 have been fully considered but

they are not persuasive.

Referring to claim 1, 19, and 20, Applicant argued DaCosta dos not teach or

suggest that *identifying a first set of attribute occurrences in the template generated*

semi-structured document using an ontology and applying the pattern within the

boundary of semi each multi-attribute data record in the template generated semi-

structured document to extract a second set of attribute occurrences as claimed in

claims 1 and 19 (Applicant's argument Page 12).

In response to this argument, it is pointed out that DaCosta in view of Gardner

teaches DaCosta **in view of** Gardner teaches the claimed limitations of claims 1, 19,

and 20. DaCosta in view of Gardner teaches identifying a first set of attribute

occurrences in the template generated semi-structured document using ontology.

DaCosta teaches identifying a first set of attribute occurrences in the template

generated semi structured document as *Considered in more detail, a snippet of relevant*

*information on a web page or other web-accessible document contains **structural,***

***contents and formatting attributes**. A salient feature of one preferred embodiment of*

the system in accordance with the present invention (referred to as WebLock)

*triangulates **on these three attributes** (structural, contents, and formatting) to find and*

lock on the target data (DaCosta, Column 6 Line 1-14*).* The method of DaCosta extracts

these attributes (Column 6 Line 14-32, *extraction instructions*) employing artificial

intelligence to perform pattern matching (DaCosta, Column 5 Line 55-67, *Artificial*

Intelligence (AI) techniques can be utilized to enable pattern matching to sure that the

relevant information will still be retrieved even if the page is modified). Gardner teaches

using as in *using an ontological approach* (Gardner Paragraph 0048). Gardner teaches

an ontology-based information management system and method wherein ontology is

used to identify a set of attributes (Gardner Paragraph 0048, Paragraph 0049, and

Paragraph 0051, i.e., *using an ontological approach*) and a boundary of each multi-

attribute data record in the template generated semi-structured document (Gardner

Paragraph 0107, and Paragraph 0017-0019, i.e*., pair wise distances between ontology*

terms measured using different distance measure). As such, the combination of

DaCosta and Gardner teaches identifying a first set of attribute occurrences in the

template generated semi-structured document using an ontology.

Additionally, DaCosta teaches the limitation *applying the pattern within the*

boundary of semi each multi-attribute data record in the template generated semi-

structured document to extract a second set of attribute occurrences as in *Artificial*

Intelligence (AI) techniques can be utilized to enable pattern matching to sure that the

*relevant information will still be retrieved even if the page is modified (*DaCosta., Column

5 Line 55-67*), triangulates on these attributes (structure, content and formatting) to find*

and lock on the target data (Column 6 Line 1-14), and *extraction instructions (DaCosta,*

Column 6 Line 14-32). As per claim 20, DaCosta in view of Gardner teaches the limitations

"an ontology for identifying a first set of attribute occurrences in the template generated semi-structured document, the ontology comprising a set of concepts and a set of attributes associated with every concept" (Gardner Paragraph 0005, i.e., *An ontology term may be a single named concept describing an object or entity*, Paragraph 0051, i.e., *An ontology may be used to enable effective syntactic and semantic mapping between any number of different entities..*);

"a boundary module for determining a boundary of each multi-attribute data record in the template generated semi-structured document" (Gardner Paragraph 0107, i.e., *pair wise distances between ontology terms measured using different distance measure* and Paragraph 0017-0019. Particularly note that measuring pair wise distances between ontology terms creates a bound of each record.); and

"a pattern module for learning a pattern for an attribute corresponding to an identified attribute occurrence of the first set in the template generated semi-structured document" (DaCosta et al., Column 5 Line 55-67, i.e. *Artificial Intelligence (AI) techniques can be utilized to enable pattern matching to sure that the relevant information will still be retrieved even if the page is modified*, Column 6 Line 1-14, i.e., *..triangulates on these attributes (structure, content and formatting) to find and lock on the target data*, and Column 6 Line 14-32, i.e., *extraction instructions*).

Therefore, DaCosta in view of Gardner teaches the limitations of claims 1, 19, and 20.

Applicant argues that *Nowhere does DaCosta teach or suggest patent matching to extract attributes based on attribute occurrences identified using an ontology* (Applicant's argument, Page 12). Again, it is pointed that DaCosta in view of Gardner teaches or suggest patent matching to extract attributes based on attribute occurrences identified using an ontology as *Artificial Intelligence (AI) techniques can be utilized to enable pattern matching to sure that the relevant information will still be retrieved even if the page is modified* (DaCosta., Column 5 Line 55-67), *triangulates on these attributes (structure, content and formatting) to find and lock on the target data* (DaCosta, Column 6 Line 1-14), *extraction instructions* (DaCosta, Column 6 Line 14-32), *An ontology term may be a single named concept describing an object or entity* (Gardner Paragraph 0005), *An ontology may be used to enable effective syntactic and semantic mapping between any number of different entities*(Gardner, Paragraph 0051).

Applicant argued that *nor an ontology for identifying a first set of attribute occurrences* (Applicant's argument, Page 12-13). In response, it is pointed out that DaCosta in view of Gardner teaches said limitation as *An ontology term may be a single named concept describing an object or entity (Gardner Paragraph 0005); An ontology may be used to enable effective syntactic and semantic mapping between any number of different entities (Gardner, Paragraph 0051); and a snippet of relevant information on a web page or other web-accessible document contains **structural, contents and formatting attributes**. A salient feature of one preferred embodiment of the system in accordance with the present invention (referred to as WebLock) triangulates **on these three attributes** (structural, contents, and formatting) to find and lock on the target data*

(DaCosta, Column 6 Line 1-14*).* Other limitation of claim 20 as taught by DaCosta in

view of Gardner can be found in the part of this office action on claim 20.

Therefore, DaCosta in view of Gardner teaches the limitations of claims 1, 19,

and 20.

Applicant additionally argues that *combined teachings of DaCosta and Gardner*

fail to teach or suggest applying two extraction techniques to a template generated

semi-structured document for extracting a first and second set of attribute occurrences

(Applicant's argument, Page 14). Please note the following teachings of DaCosta.

DaCosta discloses that *Considered in more detail, a snippet of relevant information on a*

web page or other web-accessible document contains **structural, contents and**

formatting attributes. *A salient feature of one preferred embodiment of the system in*

accordance with the present invention (referred to as WebLock) triangulates **on these**

three attributes *(structural, contents, and formatting) to find and lock on the target data*

(DaCosta, Column 6 Line 1-14*).* Additionally, the method of DaCosta extracts these

attributes (Column 6 Line 14-32, *extraction instructions*) employing artificial intelligence

to perform pattern matching (DaCosta, Column 5 Line 55-67, *Artificial Intelligence (AI)*

techniques can be utilized to enable pattern matching to sure that the relevant

information will still be retrieved even if the page is modified). Gardner teaches using as

in *using an ontological approach* (Gardner Paragraph 0048). As such, the method of

DaCosta extracts attributes in the fist place (i.e., first set of attributes), and then applies

AI techniques for pattern matching to semi-structured web documents to extract second

set of attributes.

Claim Rejections - 35 USC § 101

3.   35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

4.   Claims 1-16 and 19-22 are rejected under 35 U.S.C. 101 because the claimed

invention is directed to non-statutory subject matter.

Claims 1-16 and 19-22, especially claims 1, 19 and 20, are not statutory because

the claims do not require any physical transformation and the invention as claimed does

not produce a useful and tangible result in view of MPEP 2106 (IV)(C)(2)((B))((2))(a)

and (b)[1].

Claim Rejections - 35 USC § 103

5.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and

---

[1] MPEP (IV)(C)(2)((B))((2))(a) and (b):

For an invention to be "useful" it must satisfy the utility requirement of section 101. The USPTO's official interpretation of the utility requirement provides that the utility of an invention has to be (i) specific, (ii) substantial and (iii) credible. MPEP § 2107 and Fisher, 421 F.3d at 1372, 76 USPQ2d at 1230 (citing the Utility Guidelines with approval for interpretation of "specific" and "substantial"). In addition, when the examiner has reason to believe that the claim is not for a practical application that produces a useful result, the claim should be rejected, thus requiring the applicant to distinguish the claim from the three 35 U.S.C. 101 judicial exceptions to patentable subject matter by specifically reciting in the claim the practical application. In such cases, statements in the specification describing a practical application may not be sufficient to satisfy the requirements for section 101 with respect to the claimed invention. Likewise, a claim that can be read so broadly as to include statutory and nonstatutory subject matter must be amended to limit the claim to a practical application. In other words, if the specification discloses a practical application of a section 101 judicial exception, but the claim is broader than the disclosure such that it does not require a practical application, then the claim must be rejected.

The tangible requirement does not necessarily mean that a claim must either be tied to a particular machine or apparatus or must operate to change articles or materials to a different state or thing. However, the tangible requirement does require that the claim must recite more than a 35 U.S.C. 101 judicial exception, in that the process claim must set forth a practical application of that judicial exception to produce a real-world result.

the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

6.   Claim 1-10 and 14-22 are rejected under 35 U.S.C. 103(a) as being unpatentable over DaCosta et al. (hereinafter "DaCosta") (U.S. Patent Number 6826553) in view of Gardner (U.S. Patent Application Publication Number 2003/0177112).

Referring to claim 1, DaCosta et al. teaches is directed to a method for extracting an attribute occurrence (DaCosta et al., Figure 1, Figure 3, and Column 5 Line 37-54) from template generated semi-structured document (DaCosta et al., Column 11 Line 33-50, Column 6 Line 14-31, i.e. *stock quotes or weather data* and Column 12 Line 5-14) comprising multi-attribute data records (DaCosta et al., Column 7 Line 55 through Column 8 Line 18, i.e. *Referring first to form variations, XML encoding of a form provides key-value pairs of form parameters.*) and teaches the limitations:

"identifying a first set of attribute occurrences in the template generated semi-structured document " (DaCosta et al., Column 5 Line 55-67, i.e., *Artificial Intelligence*

(AI) techniques can be utilized to enable pattern matching to sure that the relevant

information will still be retrieved even if the page is modified; Column 6 Line 1-14, i.e.,

..triangulates on these attributes (structure, content and formatting) to find and lock on

the target data, and Column 6 Line 14-32, i.e., *extraction instructions*);

"learning a pattern for an attribute corresponding to an identified attribute

occurrence of the first set in the template generated semi-structured document"

(DaCosta et al. Column 6 Line 14-32, *capable of performing some degree of learning*);

and

"applying the pattern in the template generated semi-structured document to

extract a second set of attribute occurrences" (DaCosta et al. Column 6 Line 14-32, i.e.,

the extraction module 20 infers extraction rules and applies them to the remainder of the

data in the web page or other web-based accessible document).

DaCosta does not explicitly teach the limitations: "the boundary of teach multi-

attribute data record" and "using an ontology".

Gardner teaches the limitation:

"the boundary of teach multi-attribute data record" (Gardner Paragraph 0107, i.e.

pair wise distances between ontology terms measured using different distance measure

and Paragraph 0017-0019) and

and "using an ontology" (Gardner Paragraph 0048, i.e. *using an ontological*

approach). Gardner teaches an ontology-based information management system and

method wherein ontology is used to identify a set of attributes (Gardner Paragraph

0048, i.e. *using an ontological approach*) and a boundary of each multi-attribute data

record in the template generated semi-structured document (Gardner Paragraph 0107,

i.e. *pair wise distances between ontology terms measured using different distance

measure* and Paragraph 0017-0019).

At the time the invention was made, it would have been obvious to a

person of ordinary skill in the art to combine the ontology-based information

management system and method of Gardner with the method and system for providing

database functions for multiple Internet sources as taught by DaCosta et al. so that the

combined method and system would identify a set of attribute occurrences in template-

generated documents, determine a boundary of each multi-attribute data record in the

said documents and employ the pattern of attributes within the said boundary of each

multi-attribute data record in the template generated semi-structured document to

extract a second set of attribute occurrences. One would have been motivated to do so

in order to provide *a particularly well-developed set of semantic relationships built into

the ontology* (Gardner Paragraph 0008).

Referring to claim 2, Gardner teaches the limitation:

"the step of providing a seed ontology prior to identifying the first set of attribute

occurrences (*search results of the ontology search Gardner* in Paragraph 0089-0096

and *ontology-based search in* Paragraph 0114).

Referring to claim 3, Gardner teaches the limitation:

"the ontology is one of a seed ontology" (Paragraph 0091, i.e., *one or more*

search strings of the ontological search) "and an enriched ontology" (Paragraph 0092-

0093, i.e., *the search results of the ontological search* and Paragraph 0100).\

Referring to claim 4, Gardner teaches the limitation:

"enriching the ontology with the second set of attributes occurrences" (Paragraph

0105, i.e., *The system may, in operation 624, augment an existing ontology with newly*

discovered terms to maximize the discriminable term coverage.).

Referring to claim 5, DaCosta teaches the limiations:

"wherein the pattern is a path abstraction expression, wherein the path

abstraction expression is a regular expression that does not comprise a union operator,

and a closure operator only applies to single symbols"(Column 7 Line 29-54, i.e., *the*

recorded path", and Column 9 Line 45-52,).

Referring to claim 6, DaCosta teaches the limitations:

"wherein learning the pattern for each attribute occurrence comprises: identifying

the attribute occurrence in a data structure tree" (Column 12 Line 5-14, i.e., *the system*

infers that the user wishes to further extract information 1160 using this parental

hierarchy information); and "determining the pattern of the attribute occurrence in the

data structure tree" (Column 12 Line 5-61, i.e., *parental hierarchy information 1130-*

1140-1150 and *Using AI techniques, the system of the present invention evaluates*

parental information 1130-1140-1150 and 1170-1180-1190 to determine a most likely

pattern that the highlighted element 1150 and element 1160 cab be classified as

satisfying and then finds the next element which matches that pattern, if any).

Referring to claim 7, Gardner teaches the limitation:

"further comprising the step of generalizing the pattern of the attribute occurrence

prior to applying the pattern" (Paragraph 0105-0106, i.e., *the system may combine*

multiple available ontologies into a single validated aggregate ontology).

Referring to claim 8, DaCosta teaches the limitation:

"wherein the pattern comprises elements including a location and a format of the

attribute occurrence" (Column 6 Line 1-13, *a snippet of relevant information on a web*

page or other web-accessible document contains structural, contents and formatting

attributes).

Referring to claim 9, DaCosta teaches the limitation:

" wherein the elements are nodes in the data structure tree" (Column 12 Line 5-

42, i.e., *the lineage or ancestry tree matches"* and *Using AI techniques, the system of*

the present invention evaluates parental information 1130-1140-1150 and 1170-1180-

1190 to determine a most likely pattern that the highlighted element 1150 and element

1160 cab be classified as satisfying and then finds the next element which matches that

pattern, if any).

Referring to claim 10, DaCosta teaches the limitations:

"further comprising resolving the ambiguities in the extracted attribute

occurrences" (DaCosta et al., Column 9 Line 45-52, i.e., *extraction module* and Figure

15) comprising":

"identifying attribute occurrences in the template generated semi-structured

document matching more than one pattern" (Column 5 Line 63-67, i.e., *Artificial*

Intelligence (AI) techniques can be utilized to enable pattern matching to ensure that the

relevant information will still be retrieved. Note that it is inherent that pattern

occurrences are identified in any AI pattern matching);

"determining a pattern that uniquely matches a given attribute occurrence and no

other pattern uniquely matches the given attribute occurrence" (Column 5 Line 63-67,

i.e., *Artificial Intelligence (AI) techniques can be utilized to enable pattern matching to*

ensure that the relevant information will still be retrieved. Note that it is inherent in any

AI pattern matching that a pattern that uniquely matches a given occurrence is

determined); and

"eliminating matches between the given attribute occurrence and another pattern

that matches the given attribute occurrence and at least one other attribute occurrence"

(Column 5 Line 63-67, i.e., *Artificial Intelligence (AI) techniques can be utilized to*

enable pattern matching to ensure that the relevant information will still be retrieve").

Note that it is inherent in any AI pattern occurrences that are not unique are eliminated.

Referring to claim 14, DaCosta in view of Gardner teaches the limitations:

"providing a tree of a page" (DaCosta et al., Column 5 Line 23-29, i.e., *the*

Document Object Model (DOM) of an document and Column 12 Line 5-14, i.e., *the*

system infers that the user wishes to further extract information 1160 using this parental

hierarchy information) "and a set of attribute names of a concept of the ontology"

(DaCosta, Figure 1, Figure 3, and Column 5 Line 37-54 and Gardner Paragraph 0048,

i.e. *using an ontological approach*, Paragraph 0049, and Paragraph 0051);

"marking a node in the tree by a set of attributes present in a subtree rooted at

the node" (DaCosta, Column 12 Line 5-42, i.e. *the lineage or ancestry tree matches* and

Using AI techniques, the system of the present invention evaluates parental information

1130-1140-1150 and 1170-1180-1190 to determine a most likely pattern that the

highlighted element 1150 and element 1160 cab be classified as satisfying and then

finds the next element which matches that pattern, if any);

"determining a set of maximally marked nodes in the tree" (Gardner Paragraph

0107, i.e., *pair wise distances between ontology terms measured using different*

distance measure, Paragraph 0105, i.e. *establish a maximum set of ontological terms*,

and   Paragraph 0017-0019);

"determining a page type" (Gardner Paragraph 0093-0094 and 0089, i.e., *A*

user's background/profile is determined to determine the kind of page to be searched.);

and

"extracting a boundary according to the page type" (Gardner Paragraph 0104).

Referring to claim 15, Gardner teaches the limitation:

"wherein the page type is one of a home page and a referral page" (Gardner

Paragraph 0089). Note that the method of Gardner could identify a home page

depending on the background/profile of the user and, inherently could refer to other

pages.

Referring to claim 16, DaCosta et al. in view of Gardner teaches the limitations:

"determining a maximally marked node with a highest score among the set of

maximally marked nodes in the tree" (Gardner, Paragraph 0105, i.e. *establish a*

maximum set of ontological terms) ;

"determining whether the tree comprises a single-valued attribute" (DaCosta et

al., Column 5 Line 23-29, i.e., *the Document Object Model (DOM) of an document* ;

Gardner Paragraph 0005, i.e., *An ontology term may be a single named concept*

describing an object or entity; Gardner, Paragraph 0051, i.e., *An ontology may be used*

to enable effective syntactic and semantic mapping between any number of different

entities.... ; DaCosta Figure 9 wherein single-value/multi-value attributes are presented

in the record table; and DaCosta et al. Column 11 Line 61 through Column 12 Line 5);

"determining values of the single-marked attribute upon determining the single-

valued attribute" (DaCosta Figure 9 wherein single-value/multi-value attributes are

presented in the record table and DaCosta, Column 11 Line 61 through Column 12

Line 5);

"determining whether the tree comprises a multiple-valued attribute" (DaCosta et al., Column 5 Line 23-29, i.e. *the Document Object Model (DOM) of an document , DaCosta Figure 9 wherein single-value/multi-value attributes are presented in the record table*, and DaCosta et al. Column 11 Line 61 through Column 12 Line 5); and

"determining values of the multiple-marked attribute upon determining the multiple-valued attribute" (DaCosta Figure 9 wherein single-value/multi-value attributes are presented in the record table, and DaCosta et al. Column 11 Line 61 through Column 12 Line 5).

Claim 19 is rejected on the same basis as claim 1.

Referring to claim 20, DaCosta et al. in view of Gardner teaches the limitations:

"an ontology for identifying a first set of attribute occurrences in the template generated semi-structured document, the ontology comprising a set of concepts and a set of attributes associated with every concept" (Gardner Paragraph 0005, i.e., *An ontology term may be a single named concept describing an object or entity*, Paragraph 0051, i.e., *An ontology may be used to enable effective syntactic and semantic mapping between any number of different entities..*);

"a boundary module for determining a boundary of each multi-attribute data record in the template generated semi-structured document" (Gardner Paragraph 0107, i.e., *pair wise distances between ontology terms measured using different distance*

measure and Paragraph 0017-0019. Particularly note that measuring pair wise

distances between ontology terms creates a bound of each record.); and

"a pattern module for learning a pattern for an attribute corresponding to an

identified attribute occurrence of the first set in the template generated semi-structured

document" (DaCosta et al., Column 5 Line 55-67, i.e. *Artificial Intelligence (AI)*

techniques can be utilized to enable pattern matching to sure that the relevant

information will still be retrieved even if the page is modified, Column 6 Line 1-14, i.e.,

..triangulates on these attributes (structure, content and formatting) to find and lock on

the target data, and Column 6 Line 14-32, i.e., *extraction instructions*).

Referring to claim 21, DaCosta et al. in view of Gardner teaches the limitations:

"wherein the pattern is applied within the boundary of each multi-attribute data

record in the template generated semi-structured document to extract a second set of

attribute occurrences" (DaCosta et al. Column 6 Line 14-32, i.e., *the extraction module*

20 infers extraction rules and applies them to the remainder of the data in the web page

or other web-based accessible document and Gardner Paragraph 0107, i.e. *pair wise*

distances between ontology terms measured using different distance measure and

Paragraph 0017-0019).

Referring to claim 22, DaCosta teaches the limitation:

"wherein the database of multi-attribute data records is stored on a server

connected to the adaptive search engine application across a communications network"

(Figure 22: *Smart Server* and *Web Folders* and Column 18 Line 8-25).

7.   Claim 11 and 13 are rejected under 35 U.S.C. 103(a) as being unpatentable over

DaCosta in view of Gardner and further in view of Oommen (U.S. Patent Application

Publication Number 2003/0195890).

Referring to claim 11, DaCosta et al. in view of Gardner explicitly disclose the

limitation: "learning positive examples of the attribute and learning negative examples of

the attribute".

Oommen teaches the limitation:

"learning positive examples of the attribute and learning negative examples of the

attribute" (Paragraph 00198, Paragraph 0078-0085, and Paragraph 0116-0117, i.e., *If*

the test in block 800 returns a negative answer and *if the test in block 800 returns a*

positive answer). Oommen teaches a method and system for comparing the closeness

of a target tree to other trees wherein learning pattern comprises learning positive

examples and negative examples (Oommen, Paragraph 00198, Paragraph 0078-0085,

and Paragraph 0116-0117).

At the time the invention was made, it would have been obvious to a person of

ordinary skill in the art to combine the feature of comparing closeness of trees as taught

by Oommen to the method and system of DaCosta et al. in view of Gardner as applied

to claim 1 above so that, in the combined method and system, the learning pattern

would comprise learning positive examples of the attribute and learning negative

examples of the attribute. One would have been motivated to do so in order to provide

solution for "all string, substring and subsequence recognition algorithms" as well as " a

solution to all tree, subtree and subsequence tree recognition problems (Oommen,

Paragraph 0061).

Referring to claim 13, DaCosta et al. in view of Gardner and further in view of

Oommen as applied above with regard to claim 11 discloses the invention as claimed.

DaCosta et al. in view of Gardner and further in view of Oommen is directed to the

method for claim 1, wherein learning the pattern for an attribute corresponding to an

identified attribute occurrence of the first set in the template generated semi-structured

document comprises learning negative examples of the attribute, wherein the negative

examples are positive examples of other attributes (Oommen, Paragraph 00198). In the

method of Oommen, it is inherent that if a positive answer is returned for an occurrence,

a negative answer will be returned for a different occurrence.

8.   Claim 12 is rejected under 35 U.S.C. 103(a) as being unpatentable over DaCosta

et al. in view of Gardner and further in view of Oommen and further in view of Bruno

("Efficient Creation of Statistics over Query Expressions" by Nicolas Bruno and Surajit

Chaudhuri, ICDE 2003: Bangalore, India, March 5-8 2003,

http://www.cs.brown.edu/courses/cs227/Papers/AutoAdmin/buildsits.pdf)

Referring to claim 12, DaCosta et al. in view of Gardner and further in view

Oommen teaches the limitation:

"determining common supersequence for identified attribute occurrences

corresponding to the attribute, wherein identified attribute occurrences are positive

example of the attribute" (Oommen, Paragraph 0316, i.e., *The method of the invention*

can search for the pattern where the pattern sought is distributed over a larger

supersequence as ... and Paragraph 00198, Paragraph 0078-0085, and Paragraph

0116-0117, i.e., *If the test in block 800 returns a negative answer* and *if the test in block*

800 returns a positive answer).

DaCosta et al. in view of Gardner and further in view Oommen not explicitly

disclose the limitation: "determining a generalized supersequence by generalizing each

term in the common supersequence; and determining, for each term of the generalized

supersequence, where a term can be de-generalized".

On the other hand, Bruno teaches the limitation:

"determining a generalized supersequence by generalizing each term in the

common supersequence; and determining, for each term of the generalized

supersequence, where a term can be de-generalized" (Bruno, Page 8 Column 1 Line 6

through Column 2 Line 39).

Bruno teaches a method for generalizing shortest common supersequence

wherein terms in common supersequence are generalized. (Bruno, Page 8 Column 1

Line 6 through Column 2 Line 39). Additionally, it is inherent whether a term in a

common supersequence can be degeneralized can be determined employing this method.

At the time the invention was made, it would have been obvious to a person of ordinary skill in the art to add the feature of generalizing terms in common supersequence as taught by Bruno to the method and system of DaCosta et al. in view of Gardner and further in view of Oommen as applied to claim 11 so that the resultant method and system would comprise determining a common sequence for identified attribute occurrences, determine a generalized supersequence, and determine whether a term in the generalized supersequence can be degeneralized. One would have been motivated to do so in order to "adder the optimization problem" (Bruno, Page 6, Column 2 Line 10-13).

Conclusion

9.   A shortened statutory period for reply to this final action is set to expire THREE

MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the

shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

the advisory action. In no event, however, will the statutory period for reply expire later

than SIX MONTHS from the date of this final action.

Contact Information

10.   Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Dennis Myint whose telephone number is (571) 272-

5629. The examiner can normally be reached on 8:30AM-5:30PM Monday-Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, John Breene can be reached on (571) 272-4107. The fax phone number for

the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free).

Dennis Myint   AU-2162

JOHN BREENE
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100